

DejaVu: Intra-operative Simulation for Surgical Gesture Rehearsal

Nazim Haouchine¹, Danail Stoyanov², Frederick Roy¹, and Stephane Cotin¹

¹ Inria, Mimesis Group, Strasbourg, France,

² University College London, CMIC, London, United Kingdom

Abstract. Advances in surgical simulation and surgical augmented reality have changed the way surgeons prepare for practice and conduct medical procedures. Despite considerable interest from surgeons, the use of simulation is still predominantly confined to pre-operative training of surgical tasks and the lack of robustness of surgical augmented reality means that it is seldom used for surgical guidance. In this paper, we present *DejaVu*, a novel surgical simulation approach for intra-operative surgical gesture rehearsal. With *DejaVu* we aim at bridging the gap between pre-operative surgical simulation and crucial but not yet robust intra-operative surgical augmented reality. By exploiting intra-operative images we produce a simulation that faithfully matches the actual procedure without visual discrepancies and with an underlying physical modelling that performs real-time deformation of organs and surrounding tissues, surgeons can interact with the targeted organs through grasping, pulling or cutting to immediately rehearse their next gesture. We present results on different *in vivo* surgical procedures and demonstrate the feasibility of practical use of our system.

Keywords: Computer Assisted Interventions, Surgical Simulation, Virtual Reality, Augmented Reality

1 Introduction and Background

Surgical and interventional procedures usually require years of practice to build dexterity and instrument control skills in addition to anatomical and cognitive learning. To help surgical trainees reach a high degree of reliability and accuracy, medical simulators have been developed and significant progress has been made recently to improve their accuracy, realism and fidelity. The role of a virtual medical simulator [1] is to propose a realistic environment where procedures can be conducted and repeated in an unrestricted manner without any risk or violation of patient safety. While simulators are mostly used for training purposes [2], the last decades have also seen the use of simulation for procedure planning [3] or intra-operative assistance and guidance [4]. Numerous challenges still remain to transfer simulation technologies into enabling practice of surgical sub-tasks during the procedure itself.

Simulation for training allows task based learning of gestures and assessment of the performance of the trainee [1][2], whereas simulation for planning is meant to help clinicians in the selection of the optimal therapy by adding valuable information like tumor evolution, dissection path or risk map [3]. For the latter, patient-specific data is required

describing the organs’ geometry, physiology or tissue characteristics. To transfer planning simulation to support intra-operative assistance, the pre-operative patient-specific simulation can be used as an input and evolved during surgery to modify the anatomy to its current state taking account for physiological motion, resections and insufflation, in order to enrich the surgeon with information directly during the intervention [4]. Planning and guidance are in a sense combines for such intra-operative use often using augmented reality techniques. However, while simulation for training is now often integrated in educational curricula, its use for guidance is seldom used in operating rooms. Numerous challenges remain including: (i) the inter-patient variability of visual texture and anatomical geometry which challenge computer vision and computer graphics algorithms; (ii) patient-specific tissue characterization where parameters governing deformation are appropriate on a per case basis; and (iii) the lack of ground-truth data such as intra-operative human scans of minimally invasive surgery to validate performance and provide quality assurance.

In this paper, we present a new simulation approach, which we call *DejaVu*, that permits ”just-in-time” intra-operative simulation for surgical gesture rehearsal (see Fig. 1). This new paradigm gives the possibility for surgeons to directly build a simulation from intra-operative images and to rehearse their next actions and gestures on a patient adapted virtual environment. Using the built simulation (following Subsec 2.1), virtual interaction with organs through grasping, pulling or cutting and virtual navigation in the endoscopic scene are possible without risks for the patient. Organs deformations and attachments with surrounding tissues are computed using an underlying physical model (described in Subsec. 2.2) while final composition is generated using actual image leading to a faithful and realistic visualization (explained in Subsec. 2.3). We present compelling results in Section 3 for different surgical applications and believe this is a new effort towards bringing computational techniques to the surgeons assistance in the operating theatre.

2 Materials and Methods

2.1 Overview of a DejaVu Simulation

Our approach, illustrated in Figure 1, involves a composition function Ω that enables surgeons to virtually interact with a pre-built organ model and rehearse surgical gestures. Let \mathcal{I} be an image selected by the surgeon from the intra-operative surgical site and let \mathcal{M} be a labeled 3D mesh generated from pre-operative scans that includes the organ’s surface, internal structures such as vessels or tumors, and any surrounding anatomical information. The composition Ω permits generation of a new image \mathcal{J} that mimics physical realism in term of tissue response while maintaining visual fidelity.

The pre-operative 3D mesh \mathcal{M} allows us to build a physical model incorporating the tissue properties and biomechanical behavior. This physical model is characterized by a geometry \mathcal{M} and a stiffness matrix \mathbf{K} that computes physical properties such tissue elasticity, damping or viscosity. In general, organs are attached to their surroundings by ligaments or stiff muscles. These attachments are defined in the 3D mesh, pre-operatively as a set of fixed nodes, and lead to the binary label vector \mathbf{q} , where $\mathbf{q}(j) = 1$ means the j^{th} node is attached and $\mathbf{q}(j) = 0$ means the j^{th} node can be freely displaced.

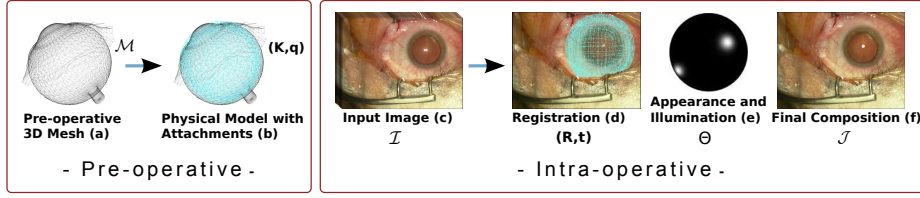


Fig. 1: Schematic illustration of DejaVu Simulation. (a) preoperative model is built from tomographic images; (b) material law, tissue properties and attachments, constitute the physical model; (c) an intra-operative image is selected; (d) 3D/2D registration is performed between the physical model in (b) and the selected frame in (c); (e) appearance and illumination are estimated corresponding to specular and diffuse components and light position; (f) the final composition is build to enable surgical gesture rehearsal.

Intra-operatively, a 3D/2D registration is manually performed because producing \mathcal{J} involves projecting the physical model onto the image. The registration computes the rotation matrix \mathbf{R} and the translation vector \mathbf{t} that relates the 3D model in world coordinates to its 2D projection in pixel coordinates. This rigid transformation is performed by the surgeon or an assistant in the operating room. Once aligned, organ appearance and scene illumination are estimated through an inverse rendering approach that estimates specular and diffuse reflection parameters and light source position. We denote Θ the set of parameters needed to produce a realistic rendering. Putting the entire process together, we can write the composition function as

$$\mathcal{J} = \Omega_{(\mathbf{R}, \mathbf{t})}(\mathcal{I}, \mathcal{M}, \mathbf{K}, \mathbf{q}, \Theta) \quad (1)$$

The output image \mathcal{J} represents an instance of a DejaVu simulation. In practice, a sequence of images is generated since a simulation implies surgeons manipulations and thus soft-tissue response and scene dynamics. Moreover, the final composition is retrieved back to surgeon's view where surgeon can virtually explore the scene, in 3D and rehearse through various interactions and visualization modes.

2.2 Organ Physical Behavior and Dynamics

To allow in situ simulation of gestures with our final composition that is displayed in the surgeon's view, our framework allows deformable model interaction. Various types of tissues can be modeled by augmenting parameters in the framework providing a range of behaviors from quasi-rigid for organs like kidneys or the uterus to hyper-elasticity for organs such as the liver [5]. The common computational pipeline is designed for spatial discretization, force/displacement computation and time discretization. Without loss of generality, we use the Finite Element Method to discretize partial differential equations of solid continuum mechanics [6]. This discretization is computed on a volumetric mesh with a finite number of degrees of freedom. This volume representation is composed of polyhedral elements and is built from a voxelization of the pre-operative 3D mesh \mathcal{M} .

Organ deformation is specified by its stress-strain relationship, which is linearized so that nodal forces \mathbf{f} can be computed from nodal displacement as: $\mathbf{f}(\mathbf{x} + \delta\mathbf{x}) =$

$\mathbf{K}(\mathbf{x})\delta\mathbf{x}$ where \mathbf{x} is a vector containing the actual position of the volume nodes, and $\delta\mathbf{x}$ their displacements. Given the relation between the position and the corresponding forces, ambient dynamics is included to capture transient events and tissue response to external event following Newton's second law to express organ motion as: $\mathbf{M} \cdot \dot{\mathbf{v}} = g(\mathbf{x}, \mathbf{v}) + \mathbf{P}$, where \mathbf{M} is the mass matrix of the organ, \mathbf{v} represents the velocities and $\dot{\mathbf{v}}$ the accelerations of the volume nodes, $g(\mathbf{x}, \mathbf{v})$ sums up forces that are related to the position or velocities of the volumes nodes and \mathbf{P} gathers external forces (such as gravity, abdominal pressure or surgical tools). This equation is often solved using time-stepping techniques [7] where time is discretized in a sequence of fixed time-steps $h = t_f - t_i$, where t_i, t_f are respectively, the time at the beginning and end of the step. The evaluation of this integration can be conducted according to various numerical schemes, however, implicit Euler is often used as it provides increased stability when dealing with large time-steps. By letting $\delta\mathbf{x} = h \cdot \mathbf{v}_f$ and $\delta\mathbf{v} = \mathbf{v}_f - \mathbf{v}_i$ we obtain the linear system of equations:

$$\underbrace{\left(\mathbf{M} - h \frac{\partial g}{\partial \mathbf{v}} - h^2 \frac{\partial g}{\partial \mathbf{x}}\right)}_{\text{Organ's mass, damping and stiffness}} \delta\mathbf{v} = \underbrace{h^2 \frac{\partial g}{\partial \mathbf{x}} \mathbf{v}_i - h(\mathbf{g}_i + \mathbf{p}_f)}_{\text{Instrument interactions}} + \underbrace{h\mathbf{H}(\mathbf{x})^T \lambda}_{\text{Organ's ligaments}} \quad (2)$$

where g_i and p_i are $g(\mathbf{x}, \mathbf{v})$ and $\mathbf{P}(t)$ at time t_i . The term $\mathbf{H}^T \lambda$ represents boundary conditions on the organ, i.e. how it is attached to its surroundings. They are modeled by enforcing some nodes of the volumetric mesh to have a null displacement following the predefined vector \mathbf{q} . \mathbf{H} is a matrix containing the constraint directions (how the nodes are constrained) while λ is a vector of Lagrange multipliers containing the constraint force intensities and is an unknown.

2.3 Organ Appearance and Scene Illumination

Visually realistic simulation requires knowledge of the organ's diffuse and specular reflection scene's illumination. Inspired by [8], we use a simplified Torrance-Sparrow reflection model that defines the specular reflection of an object's surface point as

$$\mathcal{J}_c(i) = \left[\frac{k_{d,c} \cos \theta_i}{r^2} + \frac{k_{s,c}}{r^2 \cos \theta_r} \exp \left[\frac{-\alpha^2}{2\sigma^2} \right] \right] \quad \text{with} \quad c \in \{r, g, b\} \quad (3)$$

where $\mathcal{J}(i)$ is the i^{th} image pixel value, θ_i the angle between the light source direction and the surface normal, θ_r is the angle between the viewing direction and the surface normal, α is the angle between the surface normal and the intersection of the viewing direction and the light source direction. r represents the distance between the light source and the object surface point, k_d and k_s are coefficients for the diffuse and specular reflection components respectively and include light source intensity, and σ is the surface roughness.

We want to estimate Θ that consists of the specular reflection properties (k_s, σ) , diffuse reflection k_d and light source position r from image \mathcal{I} and the registered 3D mesh \mathcal{M} . To do so, we start by directly calculating θ_r , α and θ_i using our inputs. First, the angle θ_r can be obtained using the registered geometry \mathcal{M} and camera position obtained from 3D/2D registration, then assuming a unique light source and a convex

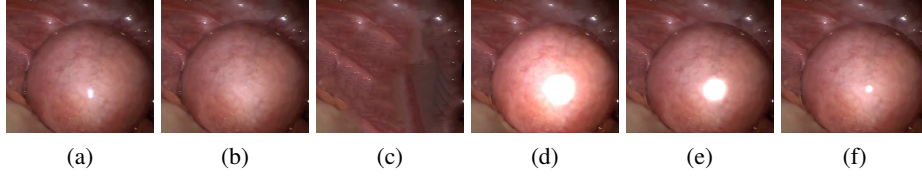


Fig. 2: Appearance and Illumination: using input image \mathcal{I} (a) diffuse image \mathcal{I}_D (b) in-painted image \mathcal{I}_B (c) and the mesh \mathcal{M} , the optimization scheme start by estimating light source position (d) then diffuse reflection (e) then specular reflection and roughness (f).

organ, light source direction can be estimated by back-projecting image specular peaks on geometry normals which permits to estimate α and θ_i . We use the method by Tan *et al.* [9] to obtain the specular regions, simultaneously we generate the diffuse (specular-free) image \mathcal{I}_D .

Assuming a Lambertian material with constant albedo, we follow a diffuse-based constraints scheme (cf Figure 2) to first estimate r knowing k_d then we refine for (k_s, σ) to finally solve for (r, k_d, k_s, σ) minimizing the squared error as

$$\underset{r, k_d, k_s, \sigma}{\operatorname{argmin}} \sum_{i \in \chi} \tau_i \left(\mathcal{I}(i) - \left[\frac{k_d \cos \theta_i}{r^2} + \frac{k_s}{r^2 \cos \theta_r} \exp \left[\frac{-\alpha^2}{2\sigma^2} \right] \right] \right)^2 \quad (4)$$

where \mathcal{I}_i the image pixel value of i and τ_i is a compensation factor used to avoid image compensation when computing the residual error. The domain χ represents the region of interest for the optimization scheme, where the diffuse image \mathcal{I}_D is used to estimate light position and diffuse reflection where the original image \mathcal{I} will be used for specular reflection estimation. Finally, once appearance and illumination have been estimated we use a ray-tracing technique to render the final pixels on a background image \mathcal{I}_B . This image is generated using inpainting technique [10] following the contour generated from the 3D/2D registration and is at the same time used to compensate revealed parts issued while manipulating the organ.

3 Results

We present results obtained from four in-vivo surgical data shown in Figure 3. These include eye surgery for the treatment of retinal pathologies, hepatic laparoscopic surgery with tumor location and resection, kidney laparoscopic surgery for partial nephrectomy and uterine surgery for localization of uterine fibroids in laparosurgery. Pre-operative 3D meshes were obtained using ITK-SNAP (www.itksnap.org) for segmentation of tomographic images. Volumetric meshes were generated using CGal (www.cgal.org) and the subsequent physical model for allowing deformable simulation was computed using the Sofa framework (www.sofa-framework.org). To present DeJaVu simulation capabilities, we select an intra-operative image from each video where we assume no instrument is present, to avoid occlusions and ease the registration and the appearance illumination estimation, in addition to the presence of specular regions that permit the direct

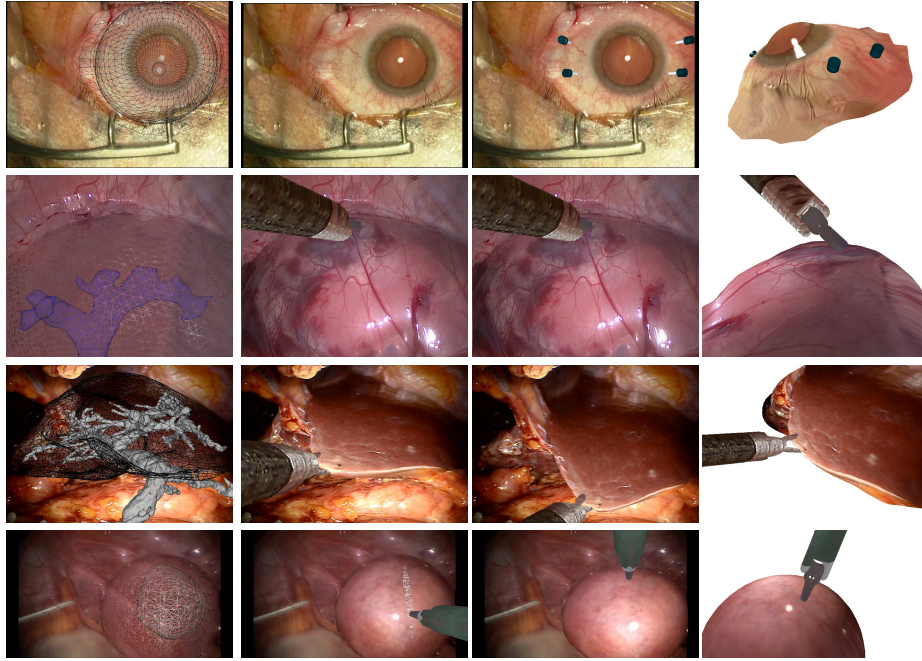


Fig. 3: DeJaVu simulation results obtained on in-vivo surgical data. From top to bottom: eye surgery, kidney surgery, liver surgery and uterine surgery. First column shows input intra-operative image with the registered pre-operative 3D mesh, second and third column show final composition with instruments interactions and last column show a 3D view of the simulation. [Scene dynamics is better seen in the additional material]

calculation of light source direction. However, surgical tools can also be easily detected and removed from the image using image inpainting, while the absence of specular blobs can be compensated with a good initialization of light source direction. The average time needed to perform the alignment is 34 seconds. The physical simulation has various parameters to be determined, depending on organ's material and characteristics: the mass m , Young's modulus E for stiffness, Poisson ratio ν for compressibility and number of polyhedral elements. For users not accustomed to using physical engines, pre-defined parameters are set according to the organ size and units and can be changed during simulation. We set the time-step $h = 0.01$ to be able to capture transient event while being computationally efficient. All simulations runs at interactive frame-rate at a minimum of 19 fps. To enable tissue manipulation by surgeons through the composite function Ω , virtual surgical instruments are added to the simulation. Surgeons can manipulate the organ in a 3D non-restricted manner: they can naturally translate and rotate the organ and the camera, perform non-rigid manipulation as stretching, torsion and compression. The framework also enables tissue/rigid contacts like grasping and

pulling and topological changes such as cutting. Moreover, bi-directional mapping is considered where the motion of the organ surface is propagated to internal structures while mechanical responses of the latter are accumulated to the whole mechanical system. Each of the surgical cases illustrated in Figure 3 depicts a surgical event or gesture

Table 1: Appearance and illumination parameters.

	light position (mm) r	diffuse reflection k_s	specular reflection k_d	roughness σ
Eye	(-0.043, 0.23, 76.65)	(0.27, 0.29, 0.30)	(0.87, 0.91, 0.95)	0.041
Kidney	(0.13, -0.45, 72.17)	(0.89, 0.99, 0.99)	(0.99, 0.97, 0.95)	0.007
Uterus	(0, 0, 87.77)	(1, 1, 1)	(1, 1, 1)	0
Liver	(-0.18, -0.001, 85.43)	(0.12, 0.11, 0.12)	(0.94, 0.93, 0.90)	0.035

where the organ needs specific modeling due to the nature of the anatomy. The results of the appearance and illumination estimation step are reported in table 1. With the *eye* experiment, the surgeon is asked to place trocars around the cornea through the conjunctiva to reach the retina located behind the conjunctiva. Tissue deformation due to the contact of the trocar with the conjunctiva is modeled as a sphere-shaped model composed of 3600 tetrahedral P1 elements derived from the sclera geometry and attached with stiff muscles to permit both rotation and elastic deformation. We used a linear co-rotated elastic model characterized by $E = 150$ kPa and $\nu = 0.45$ while mass is set to $m = 0.007$ kg. The *kidney* is modeled following a linear elastic model due to its relatively low elasticity and is built on 4219 tetrahedral P1 elements with elastic parameters $E_p = 250$ kPa and $\nu_p = 0.40$ and a mass $m = 0.115$ kg. Its vascular network represents the main source of heterogeneity and is mapped with the parenchyma and considered stiffer $E_v = 3200$ kPa and $\nu_v = 0.45$. Moreover, its suspended through its veins that represent the main ligaments. On the other hand, the *liver* is modeled as a hyper-elastic material following a Saint Venant-Kirchhoff model where its parenchyma is characterized by $E_p = 27$ kPa and $\nu_p = 0.40$. The volume is composed of 3391 tetrahedral P1 elements, and its mass is set to 1.2 kg. Similar to the kidney, hepatic and portal veins are added to the global mechanical systems and add heterogeneity and anisotropy. The vascular networks was parameterized with $E_v = 0.62$ mPa and $\nu_v = 0.45$. The ligaments are, however, more difficult to set since surrounding tissues can impact liver response depending on the intra-operative setup (abdominal pressure). Since specular regions were not accurately detected, light direction and position were manually initialized with $r = (0, 0, 100)$ directed towards the organ. The results obtained in table 1 can therefore be translated as a pure texture-mapping. Finally, the uterus is modeled as a quasi-rigid organ with small linear elasticity, restricted to small deformations and rotations around its attachments, and includes myomas visually mapped with the volume with the physical parameters: $E = 400$ kPa, $\nu = 0.35$, $m = 0.08$ kg and built on a volume of 550 tetrahedral P1 elements. Pulling and grasping are modeled by generating external forces after tool/tissue contact detection while cutting is based on re-meshing techniques.

4 Discussion and Conclusion

This paper has presented the DejaVu Simulation, a novel physics-based simulation approach for just-in-time surgical gesture rehearsal. We showed that it is possible to

obtain realistic simulation by merging intra-operative image and pre-operative tissue modeling. Our preliminary findings suggest it may be possible provide surgical assistance using computational physical models at the time of intervention. While we have demonstrated feasibility, there are limitations that need further development, such as the registration component in our framework, which needs to be able to deal with large deformation as seen in laparoscopic liver surgery from insufflation pressure. Including organ silhouettes or anatomical landmarks and integrating the surgeon efficiently in the pipeline can help constrain such complex registration. An additional challenge is to provide simulation with appropriate model parameters, where we could exploit tissue vibrations to estimate the organ's mass and stiffness to obtain patient-specific realistic physical behavior. Our work can also be extended to multiple-view images using stereoscopy or moving scope techniques will permit the modeling of the surrounding tissues and improving the appearance estimation thanks to an enriched organ texture. A user study conducted on experienced and unexperienced surgeons is obviously needed to reveal the full potential of the method while exhibiting new needs and benefits. While significant development does remain and need further work, we believe the presented framework is a promising, new step towards assistive surgical simulation in the modern operating room.

References

1. Gallagher AG, Ritter EM, C.H.e.a.: Virtual reality simulation for the operating room: Proficiency-based training as a paradigm shift in surgical skills training. *Annals of Surgery* **241**(2) (2005) 364–372
2. Aggarwal, R., Black, S., Hance, J., Darzi, A., Cheshire, N.: Virtual reality simulation training can improve inexperienced surgeons' endovascular skills. *European Journal of Vascular and Endovascular Surgery* **31**(6) (2006) 588 – 593
3. D'Albis, T., Haegelen, C., Essert, C., Fernández-Vidal, S., Lalys, F., Jannin, P.: Pydbs: an automated image processing workflow for deep brain stimulation surgery. *International Journal of Computer Assisted Radiology and Surgery* **10**(2) (2015) 117–128
4. Marescaux, J., Diana, M.: Next step in minimally invasive surgery: hybrid image-guided surgery. *Journal of Pediatric Surgery* **50**(1) (2015) 30 – 36
5. Delingette, H., Ayache, N.: Soft tissue modeling for surgery simulation. *Handbook of Numerical Analysis* **12** (2004) 453–550
6. Nealen, A., Muller, M., Keiser, R., Boxerman, E., M, C.: Physically based deformable models in computer graphics. *Computer Graphics Forum* **25**(4) 809–836
7. Anitescu, M., Potra, F.A., Stewart, D.E.: Time-stepping for three-dimensional rigid body dynamics. *Computer methods in applied mechanics and engineering* **177**(3) (1999) 183–197
8. Hara, K., Nishino, K., Ikeuchi, K.: Light source position and reflectance estimation from a single view without the distant illumination assumption. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **27**(4) (April 2005) 493–505
9. Tan, R.T., Ikeuchi, K.: Separating reflection components of textured surfaces using a single image. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **27**(2) (Feb 2005) 178–193
10. Barnes, C., Shechtman, E., Finkelstein, A., Goldman, D.: Patchmatch: a randomized correspondence algorithm for structural image editing. *ACM Transactions on Graphics-TOG* **28**(3) (2009) 24